

CLAIMS

What is claimed is:

1. A speech recognition method, the method comprising:
receiving an unknown speech signal;
generating a word graph based on an application task or based on information extracted from the unknown speech signal;
expanding the word graph into a phone graph; and
recognizing the unknown speech signal using the phone graph.
2. The method of claim 1, wherein expanding the word graph includes expanding the word graph into the phone graph based on a cross-word acoustical model.
3. The method of claim 1, wherein expanding the word graph includes:
expanding a word node of the word graph into a plurality of phone nodes for the phone graph; and
removing unnecessary phone nodes from the phone graph.
4. The method of claim 1, wherein recognizing the unknown speech signal includes recognizing known words, phrases, or sentences in the second speech signal using the phone graph.
5. The method of claim 1, wherein recognizing the unknown speech signal includes recognizing the unknown speech signal as continuous speech.
6. A speech recognition system comprising:
an input to receive an unknown speech signal; and
a processor to generate a word graph based on an application task or based on information extracted from the unknown speech signal, to expand the word graph into a phone graph, and to recognize the unknown speech signal using the phone graph.
7. The system of claim 6, wherein the processor is to expand the word graph into the phone graph based on a cross-word acoustical model.

8. The system of claim 6, wherein the processor is to expand a word node of the word graph into a plurality of phone nodes for the phone graph and to remove unnecessary phone nodes from the phone graph.
9. The system of claim 6, wherein the processor is to recognize known words, phrases, or sentences in the unknown speech signal.
10. The method of claim 1, wherein the processor is to recognize the unknown speech signal as continuous speech.
11. A machine-readable medium that provides instructions, which if executed by a processor, cause the processor to perform the operations comprising:
 - receiving an unknown speech signal;
 - generating a word graph based on an application task or based on information extracted from the unknown speech signal;
 - expanding the word graph into a phone graph; and
 - recognizing the unknown speech signal using the phone graph.
12. The machine-readable medium of claim 11, further providing instructions, which if executed by a processor, cause the processor to perform the operations of:
 - expanding the word graph into the phone graph based on a cross-word acoustical model.
13. The machine-readable medium of claim 11, further providing instructions, which if executed by a processor, cause the processor to perform the operations of:
 - expanding a word node of the word graph into a plurality of phone nodes for the phone graph; and
 - removing unnecessary phone nodes from the phone graph.
14. The machine-readable medium of claim 11, further providing instructions, which if executed by a processor, cause the processor to perform the operations of:

recognizing known words, phrases, or sentences in the unknown speech signal using the phone graph.

15. The machine-readable medium of claim 11, further providing instructions, which if executed by a processor, cause the processor to perform the operations of:
recognizing the unknown speech signal as continuous speech.
16. A speech processing method comprising:
expanding a word graph into a phone graph, the word graph including a plurality of word nodes, and the phone graph including a plurality of phone nodes.
17. The method of claim 16, wherein the expanding the word graph into the phone graph includes expanding the word graph into the phone graph based on a cross-word acoustical model or a clustered cross-word acoustical model.
18. The method of claim 17, wherein expanding the word graph into the phone graph includes:
expanding the word graph into an intermediate phone graph based on a within word acoustical model; and
expanding the intermediate phone graph into the phone graph based on the cross-word acoustical model or the clustered cross-word acoustical model.
19. A speech processing system comprising:
a memory to store a word graph, the word graph including a plurality of word nodes; and
a processor to expand the word graph into a phone graph, the phone graph including a plurality of phone nodes.
20. The system of claim 19, wherein the processor is to expand the word graph into the phone graph based on a cross-word acoustical model or a clustered cross-word acoustical model.

21. The system of claim 20, wherein the processor is to expand the word graph into an intermediate phone graph based on a within word acoustical model and is to expand the intermediate phone graph into the phone graph based on the cross-word acoustical model or the clustered cross-word acoustical model.

22. A machine-readable medium that provides instructions, which if executed by a processor, cause the processor to perform the operations comprising:

expanding a word graph into a phone graph, the word graph including a plurality of word nodes, and the phone graph including a plurality of phone nodes.

23. The machine-readable medium of claim 22, further providing instructions, which if executed by a processor, cause the processor to perform the operations of:

expanding the word graph into the phone graph based on a cross-word acoustical model or a clustered cross-word acoustical model.

24. The machine-readable medium of claim 23, further providing instructions, which if executed by a processor, cause the processor to perform the operations of:

expanding the word graph into an intermediate phone graph based on a within word acoustical model; and

expanding the intermediate phone graph into the phone graph based on the cross-word acoustical model or the clustered cross-word acoustical model.

25. A speech processing method comprising:

expanding a word graph into a first phone graph based on a within word acoustical model;

generating a second phone graph using the first phone graph based on a cross-word acoustical model;

generating new edges for the second phone graph;

removing obsolete phone nodes and edges from the second phone graph from the second phone graph; and

transferring necessary information from word end nodes to phone end nodes in the second phone graph.

26. The speech processing method of claim 25, further comprising:
merging a plurality of nodes of a same cluster in the second phone graph; and
merging edges responding the same cluster into one edge.
27. The speech processing method of claim 25, wherein the expanding the word graph into the first phone graph includes expanding the word graph into the first phone graph for all word nodes within the word graph.
28. A machine-readable medium that provides instructions, which if executed by a processor, cause the processor to perform operations comprising:
expanding a word graph into a first phone graph based on a within word acoustical model;
generating a second phone graph using the first phone graph based on a cross-word acoustical model;
generating new edges for the second phone graph;
removing obsolete phone nodes and edges from the second phone graph from the second phone graph; and
transferring necessary information from word end nodes to phone end nodes in the second phone graph.
29. The machine-readable medium of claim 28, further comprising instructions, which if executed by a process, cause the processor to perform the operations of:
merging a plurality of nodes of a same cluster in the second phone graph; and
merging edges responding the same cluster into one edge.
30. The machine-readable medium of claim 28, further comprising instructions, which if executed by a process, cause the processor to perform the operations of:
expanding the word graph into the first phone graph for all word nodes within the word graph.